COMPUTERWORLD SNIA

SNW

Driving Innovation Through the Information Infrastructure

SPRING 2011

Cloud Storage - What Is It?

A Service Provider's Perspective

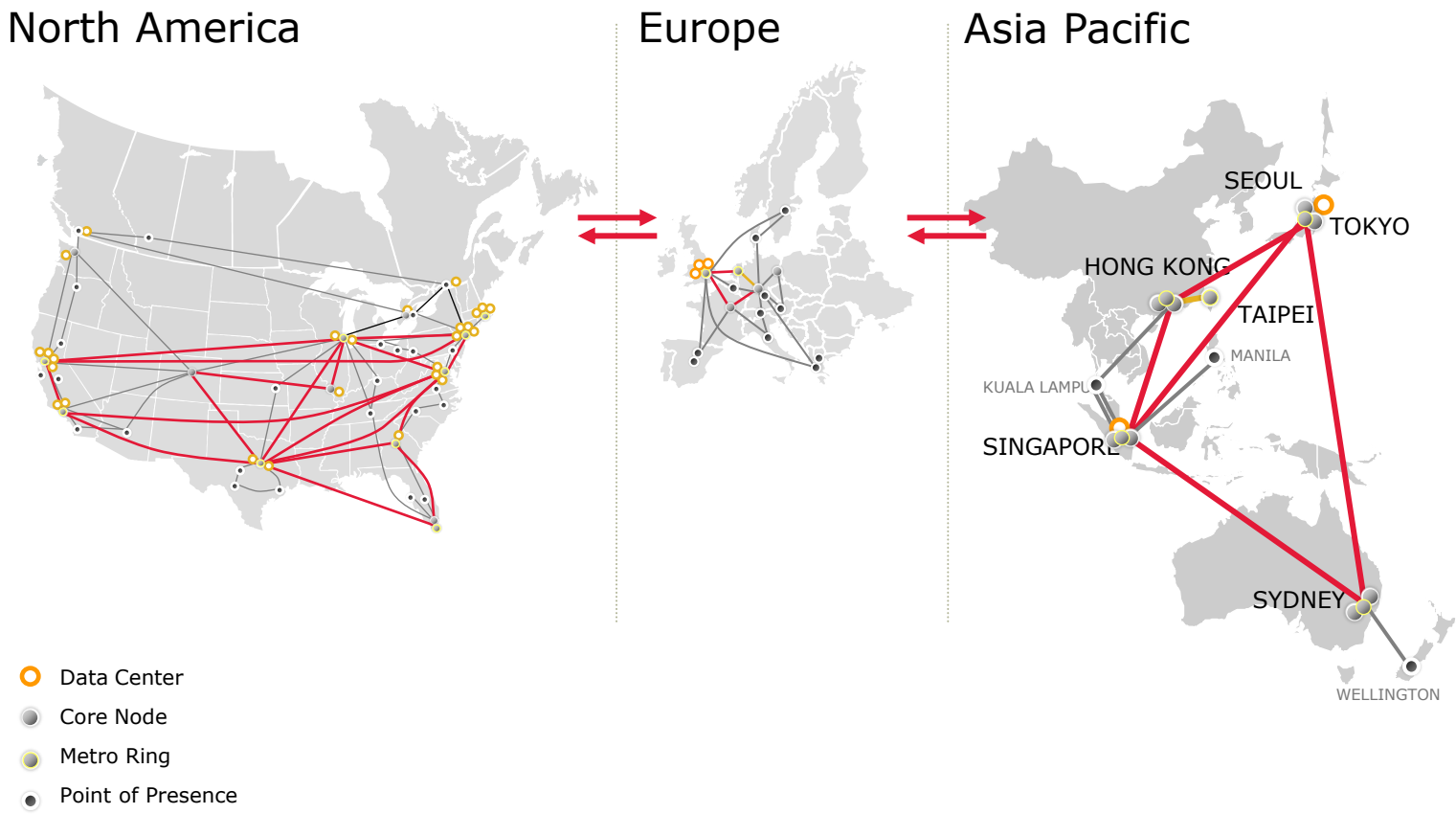
Todd Loeppke

Technical VP of Storage Architecture, Savvis

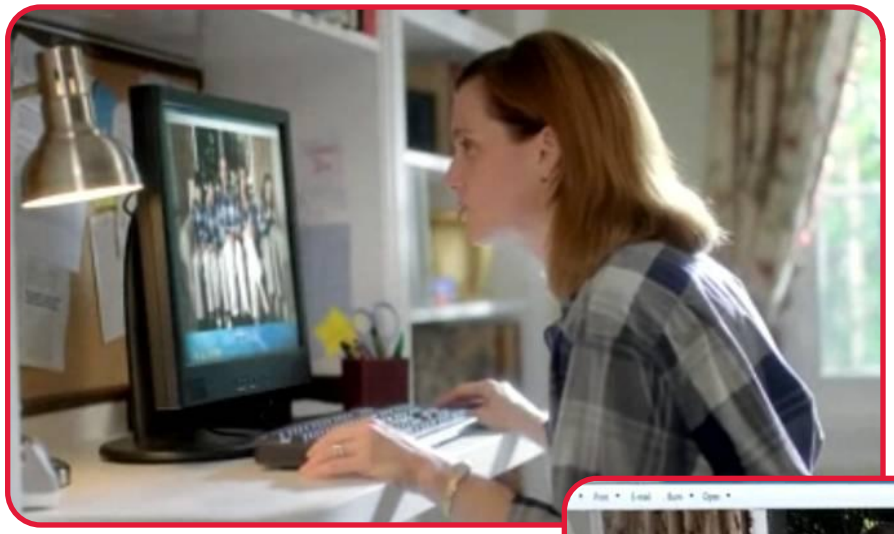
Savvis Facts

- Headquarters - St. Louis, Missouri
  - Ticker Symbol NASDAQ: SVVS
- Key Metrics
  - Nearly 2,500 unique clients, including more than 30 of the top 100 companies in the Fortune 500
  - More than 2,200 employees
  - $933 million revenue in 2010
- Services Include
  - Managed Hosting
    - Savvis Symphony Cloud Suite
    - Industry Solutions – financial, government and Software-as-a-Service (SaaS)
    - Professional Services
  - Colocation
  - Network
- The Savvis infrastructure extends to 45 countries and includes:
  - Over 22PBs of managed storage
  - 31 data centers, approximately 1.5 million square feet of raised floor space
  - 22,000 managed circuits in a private network supporting multiple application service levels
  - Tier-1 OC-192 Internet backbone with over 17,000 miles of fiber

Savvis' Global Infrastructure



If it was not already confusing enough…



“Cloud” now means…

The Internet???

Quiz: Cloud Storage is…

- ☑ A. Storage where the word “Cloud” precedes it
- ☑ B. A way for storage architects to not be cutout of the cloud buzz
- ☑ C. The third fundamental way to access storage
- ☑ D. A web service
- ☑ E. Storage available over the internet
- ☑ F. Introducing latency between now and a 10:15 presentation you are looking forward to

Agenda

- The differences between Cloud storage and Cloud computing
- What customers want
- Cloud storage vs. block and file-based storage
- Cloud storage vendor landscape
- Cloud storage service provider landscape
- Cloud storage attributes/characteristics
- Cloud storage use categories/use cases
- Cloud storage gateways/on-ramps
- Cloud storage security
- Summary

Savvis' Take on "The Cloud"

A new purchasing paradigm for managed services

Cloud computing isn't like grid or client-server

Cloud computing is like online music vs. CDs

- Usage-based billing
- Automatic delivery
- Minimal/no commitment
- Customer controlled service provisioning/ modification/termination

Not a new technology

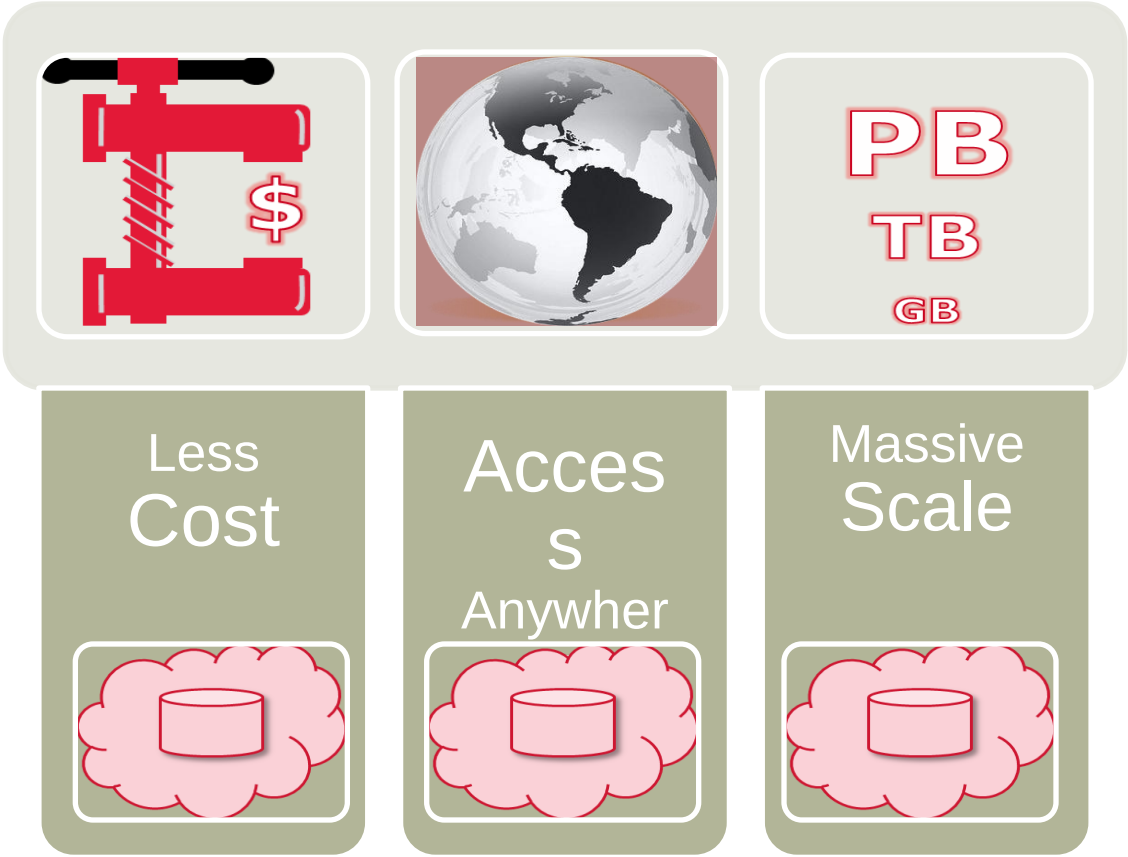
A new way to obtain data center services

Cloud Computing vs. Cloud Storage

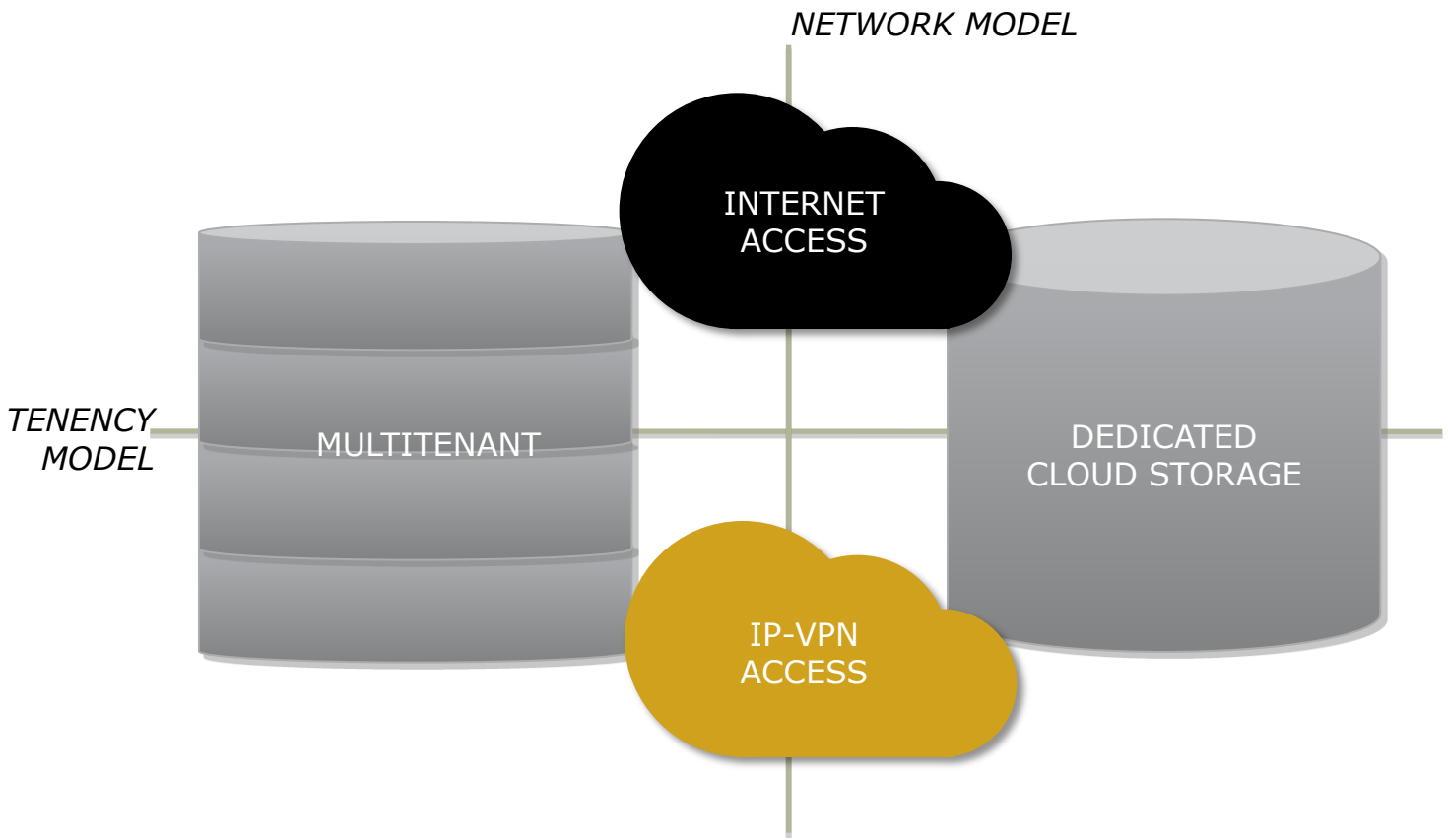
- Includes CPUs/RAM
- Hypervisor in the mix
- Can use different storage access methods
- A variety of performance requirements
- A variety of cost/GB **or** Cost/IOP
- Less flexibility in where the data “lives”

- Does not include customer facing CPUs and RAM
- Architecture may or may not include a hypervisor
- The access method is via an API (“gateways” can be layered on top)
- Lower-end performance
- Lower cost due the architecture and components
- More flexibility where the data lives

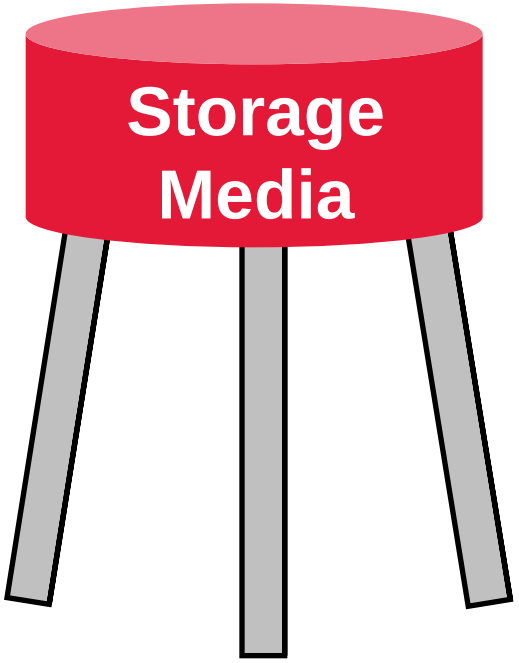
Why customers want Cloud storage



Delivery models…



The Three Fundamental Ways of Interacting with Storage



File-Based

- Filesystems
- Directories
- Files

Block

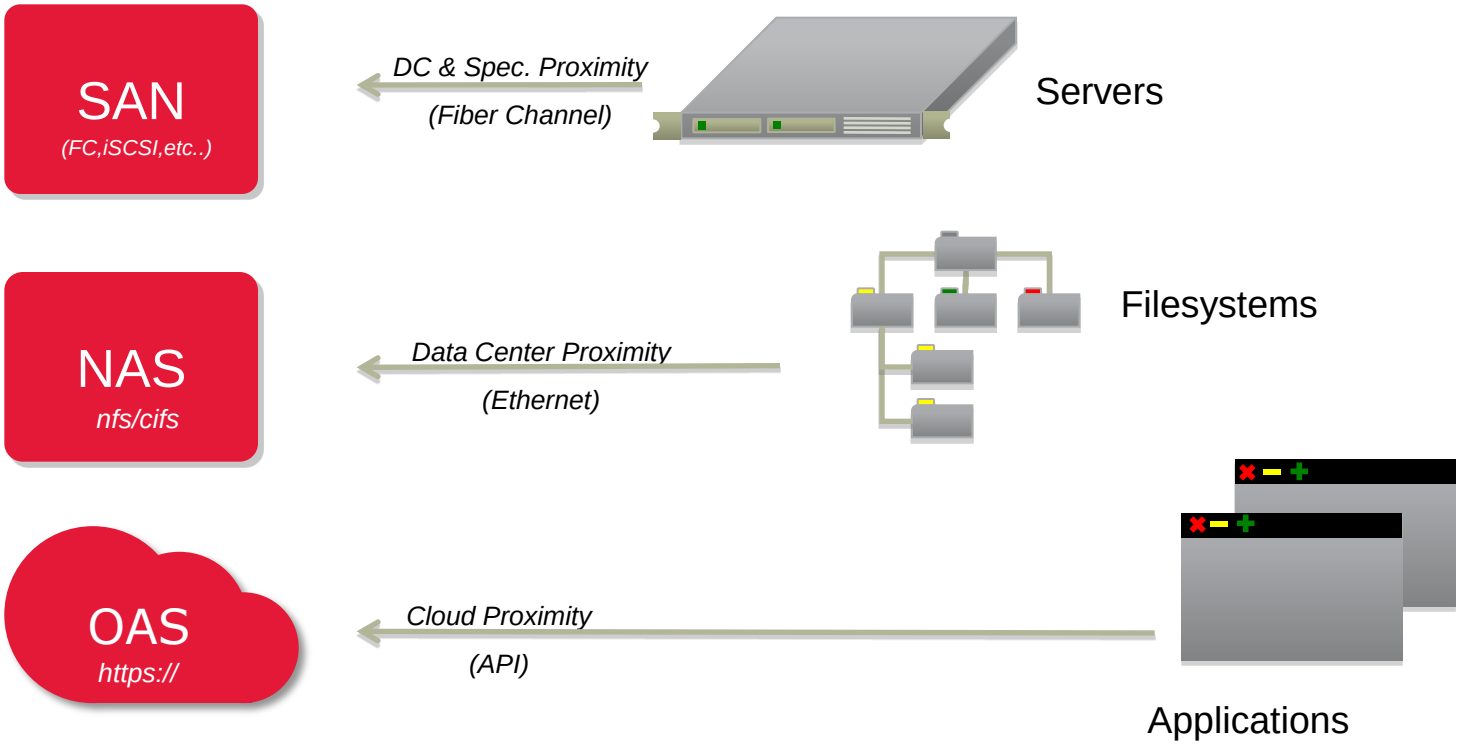
- Virtual disks/LUNs
- Volume Groups / Logical Volumes

OAS

(Object Attached Storage)

- Storage Containers
- Objects

Another Way To View It…



Cloud Storage Vendor Landscape

- Mezeo - Cloud storage platform
- EMC - Atmos
- Bycast - StorageGRID
  (Bought by NetApp May ’10)
- Cleversafe - Dispersed storage
- Amplidata - AmpliStor
- Caringo - CAStor
- DataDirect Networks (DDN) - Web Object Scaler (WOS)
- Parascale - (Bought by HDS in August ’10)
- Scality - Scality ring
- OpenStack - OpenStack object storage

Cloud Storage Service Provider Landscape

- Amazon - S3
- AT&T- Synaptic Storage-as-a-Service
- Hosted Solutions - Stratus cloud storage
- Microsoft - Azure
- Nirvanix - Storage delivery network
- Peer 1 - CloudOne storage
- Rackspace – Cloud files
- Savvis - Symphony storage cloud (currently in “alpha”)

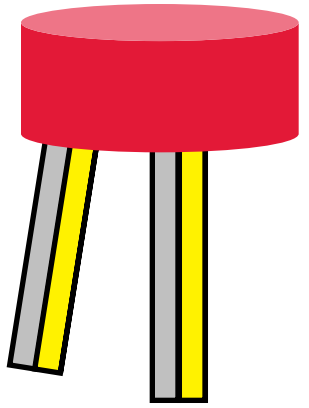
Cloud Storage Attributes/Characteristics

- Can be public or private
- Local proximity unnecessary
- Erasure coding is utilized more often than RAID
- Performance and utilization overhead are poor on “small” files/objects
- Commodity hardware is used
- Break/fix events are not a fire drill
- Usually not backed-up but may be a target for backups
- Geographic replication
- Versioning of objects used for historical protection
- The most cost affective way to store lots of data
- Starting to look more and more like a “DIY” CDN solution
- Scales better than traditional filesystems

Cloud Storage Use Categories

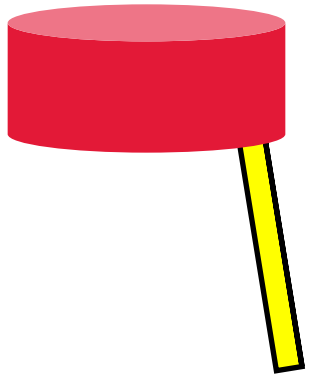
"Traditional" Storage

- ✓ NAS (file-based)
- ✓ iSCSI (block-based)



API Programming

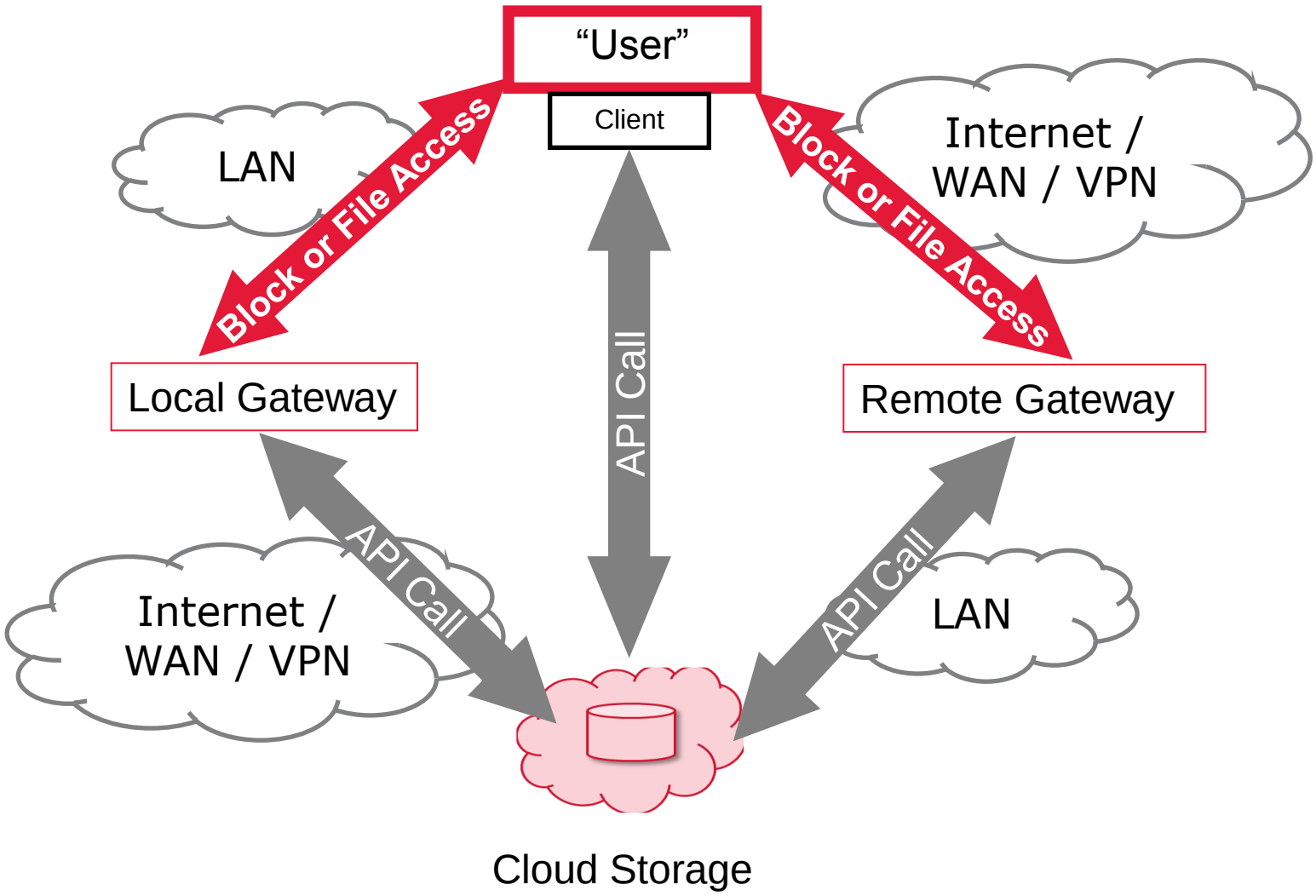
- ✓ REST/SOAP (object-based)



Cloud Storage Use Cases

- "Traditional" Storage (Using Gateways)
  - NAS (file-based)
    - ✓ Backup/archive
    - ✓ Long-tail data storage
    - ✓ Off-site tape replacement
    - ✓ Remote office file shares
    - ✓ Operational enabler – log/image transfer, image up-load
  - iSCSI (block-based)
    - ✓ Applications needing less perform-ent block storage
- API Interface
  - Content distribution
  - Media storage
  - Server BLOB storage
  - Media distribution/"DIY" CDN?
  - Data synchronization
  - Federated storage clouds

Gateways/On-Ramps



Cloud Storage Gateways
"Traditional" Storage

- Gateway/Appliances Features
  - Cache
  - Encryption
  - Compression
  - Deduplication
  - Thin provisioning
  - WORM
  - WAN optimization
  - Cloud as a tier
  - Automated tiering
  - Application Plug-ins
  - Snapshots

- Product Examples
  - Nasuni (NAS - virtual appliance - sold as a brokered service)
  - Seven10 (NAS - software)
  - Riverbed Technology (Whitewater - virtual or physical appliance)
  - Twinstrata (iSCSI - virtual or physical appliance)
  - StorSimple (iSCS - physical appliance)
  - MetaLogix (SharePoint gateway)
  - Cirtas Systems (iSCSI - physical appliance)
  - Panzura (NAS - physical appliance)
  - Gladinet (software client)
  - AtmosFox (Firefox plug-in for EMC Atmos)
  - Many more exist

API Interface

Meta Data is Key/King

Content Distribution/ Replication

THEN:

"When I was a kid… we had to rely on the storage or database administrator to replicate our data…"

NOW:

"By changing a meta data field, I can control where data lives in the storage cloud…"

API Interface

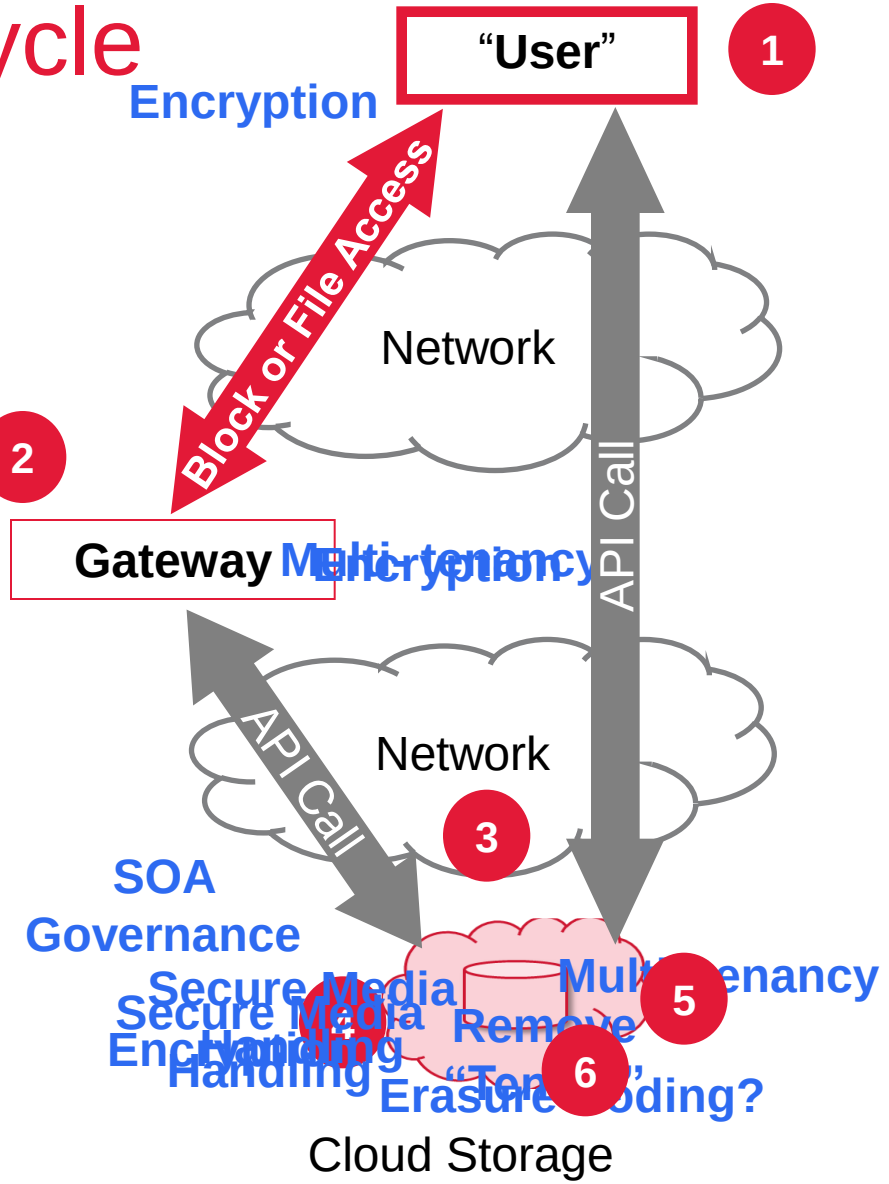
- Lack of API standards

- SNIA CDMI (Cloud Data Management Interface)
  - The Cloud Data Management Interface defines the functional interface that applications will use to create, retrieve, update and delete data elements from the Cloud.
  - SNAI Cloud storage TWG launched April, 2009
  - CDMI v 1.0 Released April, 2010

- SNIA XAM (extensible Access Method) –
  - Interface specification defines a standard access method (API) between "Consumers" (application and management software) and "Providers" (storage systems) to manage fixed content reference information storage services
  - *Working drafts produced in April, 2008*
    - *XAM - Part 1: Architecture*
    - *XAM - Part 2: C API*
    - *XAM - Part 3: JAVA API*

Security Life Cycle

1. Data creation
2. Data privacy
3. Data protection
4. Infrastructure re-use
5. Infrastructure break/fix
6. Infrastructure retirement



Summary

- Cloud storage is a useful tool for controlling storage costs
- Automation requires API programming knowledge but “gateways” can be used to augment this
- Can be leveraged to help “traditional” applications perform better and at a lower cost (Exchange/SharePoint examples)
- Can be a public or private infrastructure
- Useful for addressing large data environments
- Can provide flexible data access

Questions?

Thanks!